Occupational Downgrading of Immigrants in France

Agnès Charpin

▶ **To cite this version:**

Agnès Charpin. Occupational Downgrading of Immigrants in France. Economics and Finance. 2014. dumas-01105391

HAL Id: dumas-01105391

https://dumas.ccsd.cnrs.fr/dumas-01105391

Submitted on 20 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE
UFR 02 – SCIENCES ÉCONOMIQUES

Occupational Downgrading of Immigrants in France

Master 2 Recherche – Économie Théorique et Empirique

Présenté et soutenu par Agnès Charpin

Sous la direction de Biagio Speciale

2013/2014

Abstract

This paper analyses occupational downgrading of immigrants in France. Using the Longitudinal Survey of the Integration of First-time Arrivals (ELIPA) and the Training and Professional Qualification (FQP) survey, we show that the education-occupation mismatch of immigrants is due to their poor language proficiencies in French and their under-developed networks. Moreover, we show that immigrants are more likely to be over-educated for the occupation they hold than the native born. However, it seems that their less favourable labour market outputs converge to those of natives as their labour market experiences improve.

L'université de Paris 1 Panthéon-Sorbonne n'entend donner aucune approbation, ni désapprobation aux opinions émises dans ce mémoire ; elles doivent être considérées comme propres à leur auteur.

The University of Paris 1 Panthéon-Sorbonne neither approves nor disapproves the opinions expressed in this master thesis; they should be considered as the author's own.

TABLE OF CONTENTS

I – Introduction

1.1 Why over-education of immigrants?

There is a consensus on the key role that education holds both for individuals and the society nowadays. As a result, the proportion of graduates in the labour force has increased dramatically in almost all developed countries over the last decades. Yet, since the seventies, when the issue of education-occupation mismatch and its implications first became apparent (Freeman 1976, Rumberger 1981), many studies have documented the fact that a significant share of workers has a level of education that is not in line with their occupation. For instance, over-education of white U.S. males has been estimated to range from 11% (Verdugo and Verdugo, 1989) to 40% (Duncan and Hoffman, 1981). The literature on education-occupation mismatch lies on the idea that each occupation has a "required" education level in order for the tasks to be performed in an efficient way (Hartog, 2000). However, within an occupation, there might be workers whose actual level of education is greater or smaller than the required one: we say that there is over-education or under-education, respectively. This is what the mismatch literature aims at studying: the discrepancies that exist between actual and required levels of education, but also their extent and their effects (Chiswick and Miller, 2009).

Large scale immigration in European countries adds a new perspective to occupational downgrading: immigrants are more likely than the native labour forces to have poor proficiencies in the destination country's language, to have a poor knowledge of "how things work" in the destination country (administrative issues, how to find a decent place to live, how to look for employment, etc.), to feel bad at ease in an environment that is not theirs and, above all, to be seen as less educated than they are, because of the less-than-perfect transferability of the human capital they acquired in their home country (Chiswick and Miller, 2009). Hence, studying occupational downgrading of the immigrant population enables us to see in more details why they are less adequately matched than natives, which is an important issue in immigration economics because better knowledge of the channels through which occupational downgrading occurs can improve the support system and, as a result, benefit to both the destination country's economy and the immigrants' living conditions. Indeed, occupational downgrading is not only a "waste" of human capital, but also a cause of persisting wage penalties (Chiswick and Miller, 2008, 2011), which can aggravate inequalities between immigrants and the native born.

1.2 Mismatch in the literature: a theoretical overview

This paper contributes to a growing literature on education-occupation match of immigrants. On the technical side, empirical research on mismatch has relied on three different ways to define the "required" level of education for an occupation. The first one is the "job analyst" method. It consists in the evaluation by experts of the "required" level (and type) of education for each occupation within an occupational group. For instance, Thurow and Lucas (1972) as well as Hartog (1980) used the most sophisticated example of this type of measure: the United States Dictionary of Occupational Titles. However, this method has the disadvantage of lacking local, up-to-date information. Then, the "worker self-assessment" measure consists in the workers specifying the level of education they think is required for their job. This method was used by Duncan and Hoffman (1981), Hartog and Tsang (1987) but also by Dolton and Vignoles (2000) and Sicherman (1991). It has the disadvantage of being subjective not only to the workers' perception of their job, but also to the way the information is collected. Indeed, for instance, workers may report different levels of education depending on whether they are asked the level of education needed to *get* their job or the level of education needed to *perform* their job. Finally, the "'realized matches" method consists in observing the actual distribution of education for each occupation and considering employees who have a level of education departing from a centrality index – which can be the median, the mean (Verdugo and Verdugo, 1989) or the modal (Mendes de Oliviera and al., 2000) level of education – by more than one standard deviation as over- or under-educated.

Before giving the main results of the mismatch literature, we highlight some possible explanations for under- and over-education. First, the search and match theory suggests that the mismatch between qualifications and occupations could be due to imperfect information in the labour market. Chiswick and Miller (2009) claim that according to this view, mismatch is a temporary phenomenon because workers who take jobs for which they are over-educated can quit when they find a more favourable match and hence move up the occupational ladder (that is, limit the extent of over-education). Second, the human capital theory suggests that there exist several types of human capital (formal schooling, job experience, etc.) between which workers can trade-off. Indeed, workers could be willing to engage in occupations that require less education than they have in order to gain job experience (resulting in over-education). They could also be willing to substitute job experience for their low level of education (resulting in under-education). Alba-Ramirez (1993) confirmed this hypothesis by finding evidence on the fact that over-educated workers have less human capital such as work

experience or training. Third, Kiker, Mendes de Oliveira and Santos (2000) documented the fact that mismatch could also be due to technological changes. They argue that, in order to keep up with the pace of technological change, school-provided skills tend to be improved. Therefore, workers who acquire this type of skills and who will gradually be employed will be considered as over-educated compared to the workers who entered the workforce earlier and did not acquire these skills. Fourth, the screening hypothesis considers schooling as a way for workers to signal their unobserved ability. Because of the imperfect transferability of skills and the opacity of the quality of schooling acquired abroad, employers can have the incentive to hire immigrants that have a higher level of schooling than the "required" one in order to make sure that the tasks will be performed efficiently. Finally, some authors, such as Hartog (2000), suggest that mismatch could simply be due to differences between workers' skills and the level of skills required by the occupations offered on the labour market, and that hence, it has no particular relevance to the immigrant labour market (Chiswick and Miller, 2009).

Concerning the main results that the literature on mismatch of immigrants obtained, Aleksynska and Tritah (2013) used European data to show that immigrants are more likely to be both under- and over-educated but also that, as years of experience on the labour market increase, immigrants' labour market outcomes tend to converge to the natives' ones. They also conducted a study on selection and sorting and showed that they have a differential impact on under- and over-education. Dustmann and Frattini (2011) also used European data to highlight the general result saying that immigrants are disadvantaged relative to natives in terms of incomes, employment probability and occupational distribution. They argue that it might be due to the fact that large scale immigration is a relatively recent phenomenon in European countries and therefore that their institutions have not been sufficiently adapted to accommodate foreign born individuals yet. Chiswick and Miller (2009) used the 2000 US census to study the less-than-perfect transferability of skills. They showed that there exists an inverse relation between pre-immigration labour market experience and the likelihood of being correctly matched. They also showed that as the number of years lived in the U.S. increases, the probability of being over-educated decreases and the probability of being under-educated increases. It suggests that once the disadvantages of less-than-perfect transferability of pre-immigration skills are overcome, the positive selection of immigrants becomes realized and observable. Finally, a great share of the over-education literature deals with the returns to over- and under-education. Concerning over-education, the consensus is that over-educated

workers earn less than workers with the same educational attainment but who have occupations for which they are adequately educated; but earn more than their co-workers who have the "required" level of schooling for their job, and identical other characteristics (Kiker and al., 1997). In other words, the returns to over-education are positive but lower than to "required" education (Hartog, 2000). This added to the fact that an additional year of schooling has a lower impact on earnings for immigrants than for natives (Chiswick and Miller, 2008), makes the mismatch in the immigrant population be a particularly relevant issue.

Even though under-education is a phenomenon which extent is as important as over-education's, this paper focuses on the situation where an individual has a higher level of education than the one required to perform his job in the host country; for instance a taxi driver in the U.S. being a former scientist from Eastern Europe (Mattoo, Neagu and Özden, 2008). Moreover, we decided to use the realized matches' procedure to measure over-education of immigrants. This choice should not affect our results much. Indeed, Chiswick and Miller (2009) showed that similar findings will be drawn from analyses based on the different possible measures of over-education. Finally, and this is partly where the novelty of the paper is, we use French data on both the immigrant and native populations.

We find that occupational downgrading is more of a concern for immigrants than for the native born population. It can be mainly explained by the fact that occupational mismatch does not only depend on characteristics such as gender, age or experience in the labour market, but also on language proficiency and the extent of networks of workers. It is worth noting that the second innovation of our paper is to base our language proficiency variables not only on self-reported measures of skills but also on tests-based measures, as described thereafter.

The paper is organized as follows. In Section II, we describe the data, give descriptive evidence on over-education in France and present our variables of interest. Section III provides the results of our econometric analysis and their economic interpretation. Section IV provides some robustness checks on the definition of our over-education variable as well as our language proficiency variables, and on the type of specification used to model education-occupation mismatch. The last section concludes.

II – Data description

2.1 Datasets

In this paper we aim at explaining the reasons why we observe occupational downgrading of immigrants, more than natives, in France. This goal is quite ambitious and requires a rich, detailed and complete data set for implementation. The Longitudinal Survey of the Integration of First-time Arrivals[1] (ELIPA) was realized by the Department of Statistics, Studies and Documentation of the General Office of Immigration and Integration of the Ministry of the Interior, and in collaboration with other relevant organisms. It targets new migrants who came from countries that do not belong to the European Economic Area or Switzerland, aged 18 or more and who signed the Contrat d'accueil et d'intégration (CAI) in 2009.

The survey is longitudinal: it has been done in three waves. The first one occurred between March and June 2010 and gathers information on 6 107 immigrants. The second one took place between March and June 2011 and enabled to re-interview 78% of the first wave's population, that is 4 756 individuals. Finally, the third wave occurred between March and June 2013 and had a response rate of 77%, that is 3 573 individuals. The respondents form a representative sample of the 97 736 signatories of the CAI in 2009, but not of the whole immigrant population hosted in France, because the survey excludes:

- students, temporary workers and interns, who are considered as temporary migrants;
- workers who hold an "employee on assignment", a "skills and talents", "scientific" or "commercial" permit;
- high-level executives and their accompanying family;
- holders of a "visitor" or "artistic and cultural occupation" permit;
- war veterans and foreign patients.

These exclusions were made because ELIPA aims at analysing the integration paths of people who have received a first residence permit and who wish to settle permanently in France. As a result, it questions new residents about:

- their migration experience: motivations of migration, situation before and after migration;
- their integration in France, regarding language proficiency, professional attainment, access to housing and social life;

[1] Enquête Longitudinale sur l'Intégration des Primo-Arrivants.

- the difficulties they encountered facing administrative procedures;
- their evaluation of the CAI.

Obviously, in order to compare the levels of education of the immigrant population and the native born, we also need data on the French labour force. We found the data that we need in a database gathering information on the Training and Professional Qualification[2] survey (FQP) conducted by INSEE in order to meet the requirements in terms of planning, projection and prevision of the formation needs, and which latest version was done in 2003. This second survey gathers information on five main topics:

- professional mobility;
- initial education;
- continuing training;
- social origin;
- earnings.

It targets people aged between 18 and 65 living in an ordinary housing in France and, in its 2003 wave, contains 39 285 individuals.

2.2 Descriptive statistics

Before describing the variables we are going to use in our analysis, it seems useful to study the databases in more details. In the Appendix, Table 2 gathers the main relevant information on our basis sample. It consists of 6 107 immigrants, 50.1% of which are women and 49.9% of which are men. On average, immigrants are quite young: more than one person out of two is between 18 and 32 years-old. Moreover, the sample mostly concerns recent migration experiences: 51.9% of the individuals arrived in France less than two years before the survey was conducted. 67.7% of the respondents are married and, on average, a respondent has between one and two children. An important share of the immigrants came to France for family reasons (69.2%) but a non-negligible proportion came as refugees (13.9%) or for professional reasons (11.9%). Most of the new migrants come from Africa: 32.6% of them are from Maghreb and 27.7% of them are from Sub-Saharan Africa, which is mainly due to the colonial past of France in these areas, and their proximity to the French territory.

[2] Formation et Qualification Professionnelle.

Concerning the educational and occupational characteristics of the individuals of our two databases, in Figure 1 of the Appendix, we can see that the distribution of years of education is quite different between migrants and French workers. Indeed, up to 26% of the immigrant population has no diploma at all, whereas this proportion is only of 19.5% in the French population. Moreover, only 29.6% of the immigrants have a high-school diploma, whereas they are 42% to have one in the French population. Finally, and quite surprisingly, 26% of the immigrants have a university degree whereas they only are 19.7% to have one in the French population. Hence, in proportion, there are more highly educated people among immigrants than among the French population. Then, in the Appendix, Table 3 provides an overview of the distribution of the two populations across socio-professional categories. The fact that immigrants are over-represented in occupations that require low levels of education is striking. Only 3.1% of them hold an executive or highly intellectual position, while 13.1% of the French working population do. On the contrary, 25.6% of the immigrants are employed as unskilled workers whereas only 8.4% of the French population are.

2.3 Dependent variable

The main goal of this paper is to study occupational downgrading of immigrants, therefore we first need to construct a dependent variable stating whether there is occupational downgrading or not. To this end, we use the realized matches' procedure, which considers that employees who depart from a centrality index of years of education by more than one standard deviation are classified as over-educated for their occupation. The methodology can be decomposed in the following steps. First of all, using a variable of the ELIPA database giving the highest diploma obtained by an individual in his home country, we generate a variable equal to the number of years of education that each immigrant completed at home[3]. Second, we use the FQP dataset in order to define a "required" level of education for each occupation in France, making sure that the occupational categories are the same in the FQP and the ELIPA databases.[4] To do so, we simply calculate the mode of the highest diploma obtained by a working individual in France for each occupational group. We also save the standard deviation of years of education for each occupational group. Once we have attributed to each individual of the ELIPA database the modal value of years of education corresponding to his occupational group in France, we define the variable *match* as the difference between years of education of an immigrant and the modal value corresponding to his occupational group,

[3] See Appendix Table 1 for more details.
[4] We use the 1982 INSEE classification called Professions et categories socio-professionnelles.

accounting for one standard deviation. Finally, we define our dependent variable *overeduc* as a dummy variable which takes the value 1 if there is occupational downgrading (that is if *match* is strictly positive) and 0 otherwise.[5] It turns out that 37% of the sample's individuals can be qualified as overeducated for their occupation.

2.4 Explanatory variables

Following the existing literature on the topic and the data we highlighted in the previous section, there is evidence that the main determinants of occupational mismatch of immigrants are their language proficiency and the extent of their networks. Yet, the ELIPA database provides many variables on which we can build such indices.

Concerning language proficiency, it is considered as a key determinant in the matching process of immigrants and vacancies. Our database enables us to distinguish two types of variables: the self-reported ones and the other ones. Therefore, we decide to build two distinct dummy variables for language proficiency (*Self_report* and *Not_fluent*), the first one based on answers of immigrants to such questions as "Do you think that you speak French very well/quite well/not so well/not at all?" and the second one based on immigrants' answers to some comprehension tests done during the survey. We decided to use dummy variables stating whether the immigrant can be considered as having a very good level of French (or fluent) or not, because the distinction between an average level and a low level turned out to be insignificant. These two variables give quite different results. Indeed, it seems that immigrants tend to underestimate their language proficiency, since only 25.5% of them end up in the fluent category when we take self-reported results into account, while 49.3% of them are in this category according to the comprehension tests done during the interview. We also decide to allow for different effects of written and oral skills on occupational downgrading, hence we split *Not_fluent* into two dummy variables, *Oral* and *Written*.

Concerning networks extent, the literature on immigration economics agrees on the fact that it is a key determinant of labour market outcomes because of its effects on the economic and social integration processes. However, there is no clear consensus in the literature on the role of networks. Granovetter (1974) suggests that stronger networks can allow immigrants to have more intense social relations, which facilitates integration, whereas Lin (1995) claims

[5] It is worth noting that an alternative method would have been to compute an average of the variable *match* across individuals and use it as the dependent variable, instead of using a dummy variable. We would have obtained an index measuring the quality of matches between workers' qualifications and their occupations in the immigrant population in France.

that in the medium and long run, networks can lead to segmentation, depending on the status and varieties of the relations. We build two distinct binary variables (*Pre_mig* and *Post_mig*), which difference lies on whether the relationships were created before or after the arrival in France. It enables us to differentiate between the integration roles of both types of acquaintances. We also allow for differences between the effects of French and foreign[6] acquaintances on occupational downgrading, through the creation of two additional dummy variables (*French_network* and *Foreign_network*).

It is also well documented that experience in the labour market matters in explaining the occupational status. Hence, we compute an experience variable (*Exp*), as age minus years of schooling. Other variables affecting occupational matches are gender, age, marital status, household size, time elapsed since arrival in France, and specific characteristics of the different origin countries and of the French regions to which immigrants migrated.

Table A presents the distribution of immigrants across the two possible occupation-education match categories. We notice that the longer the time elapsed since arrival in France, the lower the share of over-educated immigrants, which suggests that there must be some convergence phenomenon of labour market outcomes of immigrants and native born as years of labour market experience increase (Aleksynska and Tritah, 2013). Moreover, there is a slightly lower share of immigrants with a very high level of French among the over-educated population, which suggests that language proficiency matters in the selection of workers by employees. Finally, there are more people who didn't know any French person before coming to France among the over-educated population than among the correctly matched or under-educated population. This suggests that French acquaintances enable migrants to integrate more easily the French labour market. This is exactly what we are going to show in the next section.

[6] By foreign acquaintances, we mean people from the same origin country as the considered migrant and living in France.

Table A: Incidence of mismatch in the French labour market, in percent.

	Overeducated	Correctly matched or undereducated
Immigrants of which:		
Report having lived in France		
Less than two years	42.6	57.4
3 – 5 years	39.1	60.9
6 – 9 years	33.8	66.2
10 years or more	24.6	75.4
Speak French fluently		
Yes	36.9	63.1
No	37.2	62.8
Had a French acquaintance before coming		
Yes	36.1	63.8
No	37.5	62.5

Source: calculations based on the ELIPA database.

III – Empirical Analysis

We estimate a binomial probit model for the probability of being over-educated versus being perfectly matched or under-educated. The model is given by:

$$Y_i = \phi(x_i'\beta)$$

Where Y_i is the probability that immigrant i is over-educated for his occupation and x_i' is the vector containing the explanatory variables outlined above. All regressions control for household size, marital status, country of origin and region of residence in France. Since we are in the probit framework, $\phi(.)$ is the cumulative distribution function of the standard normal distribution.

This section is organized as follows. First, we describe the results we obtained concerning the language proficiency variables. Second, we analyse the results about network extent. Finally, we pool our variables into a final regression and comment its predictions.

3.1 Preliminary results

Our regression results are given in Tables 4 and 5 of the Appendix. They all include the variables *Male*, *Age*, *Exp*, *Exp²*, *Length* as well as the control variables about marital status, household size, region of residence in France and country of origin.

In columns (1), (2) and (3) of Table 4, we include the variables *Self_report*, *Not_fluent* and *Oral* and *Written* respectively, in order to see which of these variables have a significant impact on the probability of being over-educated. Here, it is useful to briefly recall how *Not_fluent* was constructed. During the survey, the respondents were asked to take some comprehension tests, some of which were oral and some other written. The oral tests consisted in showing the migrant a picture, ask him or her trivial questions about it, and assess the answers in terms of oral comprehension. The written tests consisted in showing the migrant a TV program, ask him or her trivial questions about it, and assess the answers in terms of written comprehension. Then, *Not*_fluent was given the value 1 if the migrant was considered as having a very good level of French, and 0 otherwise. *Oral* and *Written* are simply a subset of this language variable, according to whether the question is about oral or written proficiency.[7]

Column (1) enables us to immediately notice that the self-reported level of French does not impact significantly the probability of being over-educated; whereas column (2) shows that the actual level of French significantly increases the probability of occupational downgrading: not being fluent in French is associated with an increase by roughly 7 points in the probability of being over-educated. Column (3) enables us to go further in this direction, by differentiating the effect of the written and oral skills on occupational downgrading. It turns out that, while having difficulties in reading and writing increases the probability of over-education by 7.4 points, the oral measure does not seem to matter significantly.

Concerning columns (4) and (5), they simply pool together the self-reported and the test-based variables together in order to confirm our previous results.[8] Again, we see that the self-reported measure do not explain over-education significantly whereas the test-based measures do.

[7] Obviously, this methodology has important drawbacks, which will be discussed in Section IV.

[8] Note here that *Not_fluent* and *Oral* and *Written* cannot be used in a same regression: since they were constructed using the same variables, it would lead to colinearity and therefore to biased coefficients.

In Table 5, we investigate how having networks in France affects the immigrants' labour market outcomes. To conduct our analysis, in column (1), we first decide to distinguish networks pre migration and networks post migration to see whether knowing someone at the time of arrival makes a difference in terms of education-occupation matching or not, using *Pre_mig* and *Post_mig*. It turns out that both variables are statistically non-significant. It can be explained by the fact that in these variables, we included any type of acquaintances without differentiating between, for instance, having some relatives living in France and knowing a French person. Yet, it is reasonable to assume that these two types of acquaintances actually have some very different potential effects. It is very likely that someone coming from the migrant's country is not as integrated as a French person, and therefore cannot provide the same type of help in the integration process. To account for this difference, we further decide to distinguish both types of acquaintances through the variables *French_network* and *Foreign_network*, and column (2) provides our findings. We immediately notice that, as expected, what seems to matter is not whether you had acquaintances in France before or after your arrival, but rather their nationality. Knowing someone of French origin before moving to France has a significant and negative effect on the probability of being over-educated whereas knowing someone from the same origin country (friend, family) does not.

3.2 Results and interpretations

We pool our main variables in a single regression, which results are given in Table B. Our first result is that not being fluent in French increases the probability of being over-educated by 6.4 points[9]. The underlying explanation is that fluency can be seen by employers as a key condition for efficiency, since it widely improves communication between employees and therefore, possibly, productivity. Hence, an employer will tend to prefer hiring an immigrant worker fluent in French or a native worker rather than an immigrant worker whose language proficiency is limited, for a given level of education. Therefore, those workers whose language proficiency is limited will be constrained to apply for jobs for which they are over-educated but where they have a greater chance to be hired. To go further in this direction, our results show that not having good written skills increases the probability of being over-educated by 7.1 points, while the oral skills do not seem to matter much (c.f. column (2)). This result was not expected. Indeed, in the selection process of their employees, it would

[9] Note that the fact that the coefficients on the language proficiency variables could be biased due to the close link between an immigrant's level of French and his or her country of origin has been taken into account through the introduction of country-of-origin dummy variables.

Table B: individual determinants of occupational downgrading

	Dependent variable: overeduc	
	(1)	(2)
Not_fluent	.0641**	
	(.0318)	
Oral		.0192
		(.0382)
Written		.0708**
		(.0358)
French_network	-.0783**	-.0789**
	(.0358)	(.0349)
Foreign_network	-.0036	-.0001
	(.0544)	(.0539)
Male	-.1589***	-.1609***
	(.0438)	(.0437)
Age	.0592***	.0594***
	(.0042)	(.0042)
Exp	-.0432***	-.0425***
	(.0098)	(.0098)
Exp^2	-.0003*	-.0003*
	(.0002)	(.0002)
Length	.0063	.0060
	(.0050)	(.0050)
Observations	825	825
Adjusted R^2	0.3368	0.3391

Notes: This table reports the results of our probit regressions. The data are taken from the first wave of The Longitudinal Survey of the Integration of First-time Arrivals (ELIPA). Dummy dependent variable: individual education-occupation match category, taking the value 1 if the individual is over-educated. All regressions control for marital status, country of origin, household size and whether the migrant lives in IDF or not. The regressions present ROBUST standard errors. *, **, and *** indicate significance at the 10%, 5% and 1% level, respectively.

seem more logical that employers base their willingness-to-hire on the oral skills of a worker, since it is what is the easier to assess during a job interview. One possible explanation is that written skills are more discriminant than oral skills in our sample: 69.4% of the respondents

answer all the oral questions properly whereas only 39.4% of them give a correct answer to all the written questions. It highlights the fact that a great majority of immigrants have sufficient oral skills to be considered as having a good level of French whereas it is more difficult for them to acquire high written skills. This is particularly striking when looking at Figure A. To understand this Figure, it is worth explaining how we constructed the language proficiency variables. *Oral* and *Written* were constructed using immigrants' answers to some comprehension tests, which were sorted in the following categories:

Oral questions:

- Understands the question and gives a correct answer;
- Understands the question but gives an answer in his or her own language;
- Understands the question but gives a wrong answer in French;
- Does not understand/does not answer the question.

Written questions:

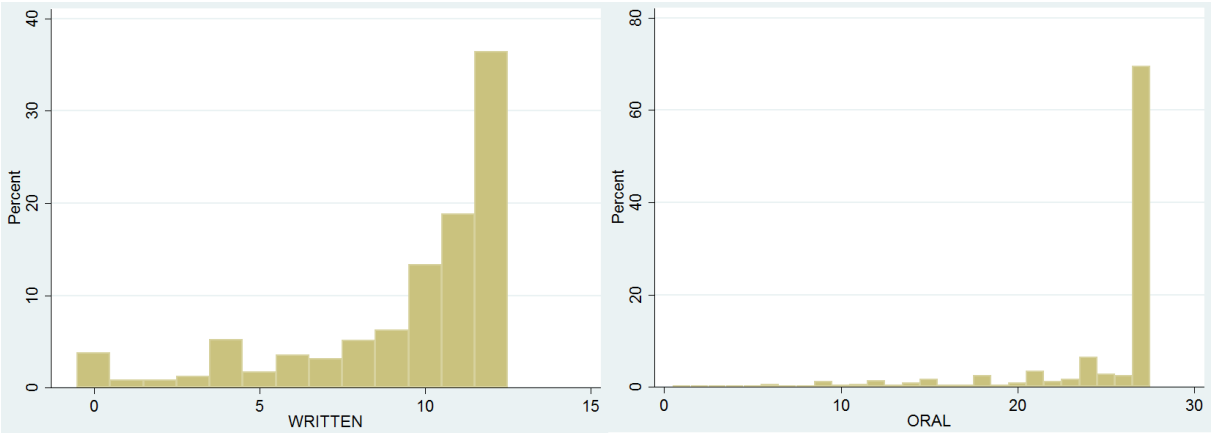
- Correct answer (including mispronunciations);
- Understands the question but gives a wrong answer;
- Reads the question but does not understand it;
- Does not answer.

To create our dummy variables, we first attributed to each answer a number of points: 3 if the migrant answers correctly, 2 if he or she understands the question but gives a wrong answer, 1 if the migrant understands the question but answers badly or if he or she cannot understand what she reads, and finally 0 if the migrant does not answer the question. Then, we sum all these variables into two variables: one stating the oral level of French (called *ORAL*) and the other one stating the written level (called *WRITTEN*). Finally, we had to choose the thresholds above which we can consider that immigrants have a high level of French (or are fluent). To do so, we computed the median values of the variables[10], and found them equal to 27 for the oral one and 11 for the written one. Therefore, we defined that people with a level of French such that they obtained a score strictly lower than 27 in oral proficiency will have a value of 1 for the variable *Oral* and those who obtained a score of strictly lower than 11 in written proficiency will have a value 1 for the variable *Written*. *Not_fluent* was constructed following

[10] *ORAL* ranges between 0 and 27 whereas *WRITTEN* ranges between 0 and 12.

the same methodology but without distinguishing both types of variables. Figure A shows the distribution of individuals across our *ORAL* and *WRITTEN* variables.

Figure A: distribution of individuals across the language proficiency indexes



Source: ELIPA, wave 1.

We can see that while the immigrant population is quite heterogeneous in terms of written skills in French, a large majority of them have the highest level of oral French allowed by the method we use. This is why our results must be taken carefully: the comprehension tests done during the survey are very trivial and it is likely that they lead to an over-estimation of the immigrants' proficiency in French (predominantly for oral skills, as shown in Figure A), whatever the methodology used to interpret them. For instance, one of the oral comprehension tests consists in showing the respondent a picture of a house and ask him or her "What colour are the walls of the house?"; yet, answer correctly this question does not require a fluent level of French.

Then, both regressions show that having a French acquaintance before moving to France decreases the probability of occupational downgrading by roughly 7.8 points whereas having a friend or a family member living in France does not have a significant impact on over-education. These results confirm the idea that a French citizen is more able to help a foreign worker to settle down in France, since he or she is more familiar with the French housing system, the French labour market and is also more likely to have a wide network of acquaintances in France than an immigrant. This is supported by De Palo, Faini and Venturini (2006), who find that, even after controlling for immigrants' individual characteristics, immigrants tend to socialize less than natives.

We also find that gender is a main determinant of over-education: being a male decreases the probability of being over-educated by roughly 16 points, compared to being a female. This result is puzzling, since it goes in the opposite direction than what was found, for instance, by Aleksynska and Tritah (2013) or Kiker and al. (1997). On the one hand, a possible explanation is the fact that in our sample, there are roughly 50% of women and that most of them did not come to France to improve their occupational status but for family reasons: as shown in Table 2 of the Appendix, 80.3% of the women in the sample came for family reasons (against 58.1% of the men) while only 4.7% of them came for professional reasons (against 19.2% of the men). As a result, it is possible that they tend to take jobs for which they are over-qualified more easily than men, which could result in a sample bias. On the other hand, another possible explanation is the discrimination of women in the French labour market. Indeed, after computing the share of over-educated women in the French labour force as well as the share of over-educated men using the FQP database, we found that while 14.8% of French women are over-educated, only 7.4% of French men are. As a result, we can think that we find a positive effect of being an immigrant woman, compared to an immigrant man, on the probability of being over-educated not because immigrant women tend to be more over-educated than immigrant men but because in general, women tend to be more over-educated than men in France.

Then, we notice that the probability of being over-educated increases significantly with age. It can be due to the fact that older migrants have a shorter payoff period in the destination country, which creates a disincentive to look for better matches because it is time-consuming, and can induce them to accept occupations for which they are over-educated more easily.

Another important result is that experience has a significant and negative impact on the probability of being over-educated: an extra year of experience decreases the probability of occupational downgrading by about 4.3 points. This result is supported by the search and match theory, or on-the-job search theory, which says that individuals accept jobs with lower educational prerequisites as an investment into the labour market and in order to have better career prospects for the future.

Finally, we included the variable *Length* because the theory suggests that as time elapses, immigrants tend to integrate the society of their destination country (economically and socially), which is likely to have a negative impact on the probability of being over-educated. As Table B shows, our specification does not enable us to observe this channel through which

over-education is impacted. It is probably due to the inclusion of the language proficiency and networks variables, since once we account for them, *Length* captures the only effect of having lived in France for a certain number of years. However, when using the second wave of ELIPA[11] and computing the share of over-educated people following the same methodology as before, we find that the proportion of workers who are occupationally downgraded decreased from 37% in 2010 to 35.6% in 2011, which, in line with the existing literature's results, shows that a longer time spent in France is associated with better labour market outcomes.

IV – Robustness checks

In this section, we perform a series of robustness checks involving both the definition of our dependent and independent variables, and the econometric methodology we have implemented in our main analysis.

4.1 On the dependent variable

To measure occupational downgrading in the immigrant population, we built an index measuring the quality of the match between workers' qualifications and their occupations. Adopting the realized matches' procedure, we considered educational attainment's distribution for each occupation and stated that employees who depart from a centrality index by at least one standard deviation are classified as over-educated. In this paper, we decided to use the mode of educational attainment for each occupation as the level of education required to perform a job efficiently, following Casarico, Facchini and Frattini (2012). Nevertheless, it seems fair to undertake some robustness checks on this choice. To this end, we decide to use the same methodology as before, but taking successively the median and the mean value of educational attainment as the "required" level of education for each occupational group. The results are reported in Table 6 of the Appendix. Column (A) shows our baseline results, while column (B) and (C) respectively show the results obtained when using the mean and median values of educational attainment.

We immediately notice that the impacts of our variables of interest vary across regressions. Indeed, the positive effect on over-education of having poor written skills loses part of its

[11] The second wave was conducted one year after the first wave and it enabled to re-interview 78% of the baseline sample's individuals.

extent when using the alternative measures for over-education. It goes from 7 to around 1 percentage point, and even loses significance in column (B). Similarly, the negative effect of having a French network on occupational downgrading goes from almost 8 to 1.8 in column (B), and to less than 1 percentage point in column (C). Following a similar pattern, all variables see their incidence on the probability of over-education reduced. In column (C), we even obtain a slightly positive and significant impact of being a male on the probability of being over-educated, which goes in the opposite direction than our baseline result. As argued by Mendes de Oliveira, Santos, and Kiker (2000), these different results can be explained by the fact that the mode is a more accurate centrality index than the median or the mean in the case on a discrete distribution (such as the distribution of years of education) because it is less sensitive to the presence of outliers in the data. They illustrate this idea using a numerical example, which we adapt to our purpose. Let's consider a group of ten workers of the same occupational group, of which eight are high-school graduates (12 years of education), the ninth has completed the first three years of university degree (15 years), and let *x* denote years of schooling of the tenth worker. According to the mode criterion, the ninth worker will be considered as over-educated no matter what level of education the last worker has. In contrast, according to the mean-centered criterion, the ninth worker will be classified as over-educated only for specific values of *x*. One can think that, as soon as within an occupation, a large majority of workers has a certain level of schooling, it can be considered as the required level for the job, and that over-education of workers within this occupation should not depend on the level of education of the outliers. As a result, in their paper, they decide to keep in their database only the individuals working in occupations which modal level of education is shared by at least 60% of workers in that occupation. This method ensures the minimization of measurement errors. Yet, in our case, there is no occupational group which has its modal value shared by more than 60% of its population. The proportion of people sharing a modal value ranges from 30.9% to 52.6% in our database. Hence, because there is not a sufficient majority of people in our modal categories, there is no proper way to apply this method to our study and we suspect that this is why we obtain results that differ from our baseline results.

4.2 On the explanatory variables

As noted in Section III, there exist several drawbacks to the methodology used to create our language proficiency variables. *Not_fluent*, and therefore *Oral* and *Written* were constructed using immigrants' answers to some comprehension tests, following the methodology explained in Section 3.2.

The results that we obtained using this method are very subjective to the choices we made along the variables' construction process. Hence, we decided to undertake a robustness check on their definition. To this end, we constructed an alternative indicator for the overall level of French, using the modal value, instead of the median one, as the threshold above which we consider a high level of French. The results are reported in Table 7 of the Appendix. Columns (A) and (B) show the regression results in which we use the base indicators for language proficiency, while column (A') and (B') respectively show the results obtained when using our alternative measures of language proficiency, that is the ones using the mode language level as the threshold, instead of the median. Comparing columns (A) and (A'), we can see that our results are very similar across regressions: the coefficients we obtain in (A') are insignificantly different from the ones in (A). Similarly, regressions (B) and (B') give very close results. Therefore, we can say that our results are robust to the method used to measure language proficiency.

4.3 On the econometric methodology

Finally, we have assessed the robustness of our findings to the choice of an alternative econometric methodology. Rather than using a probit specification, we decided to use a linear probability model (LPM). It amounts to ignoring that the dependent variable is a dummy variable and it has some disadvantages, such as the fact that it can predict probabilities greater than 1 and smaller than 0, and the fact that its leads to non-normal and heteroscedastic error terms, and this is why we did not use it as our baseline specification. Nevertheless, it enables to check the validity of our results.

Table 8 of the Appendix reports our findings. Columns Probit (1) and LPM (1) report the results of the regressions in which we use *Not_fluent* whereas columns Probit (2) and LPM (2) report those in which we use *Oral* and *Written*. We notice that in the LPM specifications, we obtain smaller coefficients in absolute value than in the probit ones. However, these differences are not significant: what matters is that our results keep the same sign and significance level than our benchmark regressions' ones. This broadly confirms the results we have obtained with the probit specifications. Having a poor level of French significantly increases the probability of being over-educated, while having French acquaintances before arriving in France decreases it significantly.

VI – Conclusions

In this paper, we use the ELIPA (Ministry of the Interior) and FQP (INSEE) databases to study the main determinants of occupational downgrading of immigrants in France. First, we show that occupational downgrading is mainly determined by the immigrants' language proficiency and the extent to which they had acquaintances living in France before arriving, but also depends on individual characteristics such as gender and age. It explains why the share of over-educated workers is significantly lower in the French labour force (11.2%[12]) than in the immigrant population (37%). In terms of policy implications, it suggests that efforts should be exerted to improve immigrants' skills in French in order to limit the extent of occupational downgrading. In the first wave of ELIPA, less than one quarter of immigrants say they have been offered to follow French lessons in the context of the reception system of immigrants, following the signature of the CAI. Three years later, 54% of the respondents of the third wave had finished the training and, among those who finished it between waves 2 and 3, 71% claimed that it is not sufficient to reach a good level in French. As a result, we can say that there is room for improvements in the design of these French lessons and their implementation. Second, we find not only that the probability of occupational downgrading of immigrants significantly decreases with experience, but also that, when using the second wave of ELIPA, the share of over-educated workers of our sample decreased from 37% in 2010 to 35.6% in 2011. It stresses the well-known phenomenon of convergence of the immigrants' labour market outcomes to those of the native born as time elapses, due to a gradual integration process in the French society.

In light of our results, it can be asserted that having a level of schooling which is in line with the "required" level is not a sufficient condition to get a job. Also, our results suggest that over-education is a rather short- or medium-term problem in the working lives of immigrants in France. As their level of French improves and they develop their networks, but also as they gain experience in the labour market, the immigrants' integration is facilitated and, as a result, their occupational outcomes improve. Hence, our findings go in the same direction than the theory saying that over-education might be the result of a trade-off between formal education and on-the-job training. As a result, it could be interesting to go further in this direction and perform a more accurate study on this topic using the data on formal schooling and on-the-job training that can be found in the three waves of ELIPA, in order to see to which extent the two

[12] Data computed from the INSEE's FQP database.

different explanations impact over-education. Another possible interesting extension of this paper would be to estimate the impact of occupational downgrading on immigrants' earnings in France using the ELIPA database, and compare it to the effect of over-education of natives on their earnings thanks to the FQP database. Finally, further research could focus not on the probability of being over-educated but on modelling the time spent over-educated by immigrants and compare it to the natives'.

VII – References

Alba-Ramirez, A. (1993), Mismatch in the Spanish Labor Market: Overducation? *The Journal of Human resources*, 28(2), 259-278.

Aleksynska, M. and Tritah, A. (2013), Occupation-education mismatch of immigrant workers in Europe: Context and policies, *Economics of Education Review*, 36, 229-244.

Casarico, A., Facchini, G. and Frattini, T. (2012), What Drives Immigration Amnesties? *CESifo Working Paper, No. 3981*.

Chevalier, A. (2003), Measuring Over-Education. *Economica*, New Series, 70(279), 509-531.

Chiswick, Barry R. and Miller, Paul W. (2009), The International Transferability of Immigrants' Human Capital Skills. *Economics of Education Review*, 28 (2), 162-169.

Chiswick, Barry R. and Miller, Paul W. (2008), Why is the payoff to schooling smaller for immigrants? *Labour Economics*, 15(6), 1317-1340.

Chiswick, Barry R. and Miller, Paul W. (2009), Does the Choice of Reference Levels of Education Matter in the ORU Earnings Equation? *Economics of Education Review*, 29(6), 1076-1085.

Chiswick, Barry R. and Miller, Paul W. (2011), Educational mismatch: are high skilled immigrants really working at high-skilled jobs and the price they pay if they aren't? In B. Chiswick (Ed.), *High skilled immigration in a globalized labor market*. Washington DC: American Enterprise Institute.

De Palo, D., Faini, R. and Venturini, A. (2006), The Social Assimilation of Immigrants. *IZA Discussion Paper* No. 2439.

Dolton, P. and Vignoles, A., (2000), The incidence and effects of overeducation in the U.K. graduate labour market. *Economics of Education Review*, 19, 179-198.

Duncan, Greg J. and Hoffman, Saul D., (1981), The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1), 75-86.

Dustmann, C. and Frattini, T. (2011), Immigration: The European Experience. *Centro Studi Luca d'Agliano Development Studies Working Papers, No. 326*.

Dustmann, C. and van Soest, A., (2001), Language Fluency And Earnings: Estimation With Misclassified Language Indicators. *The Review of Economics and Statistics*, MIT Press, 83(4), 663-674.

Freeman, Richard B. (1976), The overeducated American. New York: *Academic Press*.

Granovetter, Mark S. (1974), Getting a job: A study of contacts and careers. *Harvard University Press*.

Hartog, J. (1980), Earnings and capability requirements. *Review of Economics and Statistics*, 62(2), 230-240.

Hartog, J. (2000), Over-education and earnings: Where are we, where should we go? *Economics of Education Review*, 19(2), 131-147.

Hartog, J., Tsang, M., (1987), Estimating, testing and applying a comparative advantage earnings function for the US 1969-1973-1977. *Research Memorandum* 8709, Universiteit van Amsterdam, Department of Economics.

Lin, N., (1995), Les ressources sociales : une théorie du capital social. *Revue Française de Sociologie*, 36(4), 685-704.

Mattoo, A., Neagu, I.C. and Özden, C. (2005), Brain Waste? Educated Immigrants in the U.S. Labor Market. *World Bank Policy Research Working Paper* 3581.

Mendes de Oliveira, M., Santos, M.C. and Kiker, B.F. (2000), The role of human capital and technological change in overeducation. *Economics of Education Review*, 19, 199-206.

Rumberger, R. (1981), The rising incidence of overeducation in the US labor market. *Economics of Education Review*, 1(3), 293-314.

Sicherman, N. (1991), Overeducation in the labor market. *Journal of Labor Economics*, 9(2), 101–122.

Thurow, L.C. and Lucas, R.E.B., (1972), The American distribution of income: a structural problem. A study for the Joint Economic Committee, US Congress, *Government Printing Office*, Washington DC.

Verdugo, R. and Verdugo, N. (1989), The impact of surplus schooling on earnings: Some additional findings. *Economics of Education Review*, 22(4), 690-695.

VIII – Appendix

Table 1: Compiling the level of education data

The level of education reached by immigrants is formed from the question "What is the highest degree that you have completed at the end of your initial formation?". The categorical data were converted into a discrete variable using the following scheme.

Education category	Assumed level of education
No diploma	0
Primary school certificate or equivalent level foreign certificate	5
Certificate of general education, certificate of elementary education or equivalent	9
CAP, BEP or equivalent	11
Technical or vocational baccalaureate or equivalent	12
General baccalaureate, higher national certificate, legal assistant, DAEU	12
Undergraduate university degree	15
Masters or postgraduate degree and higher	17

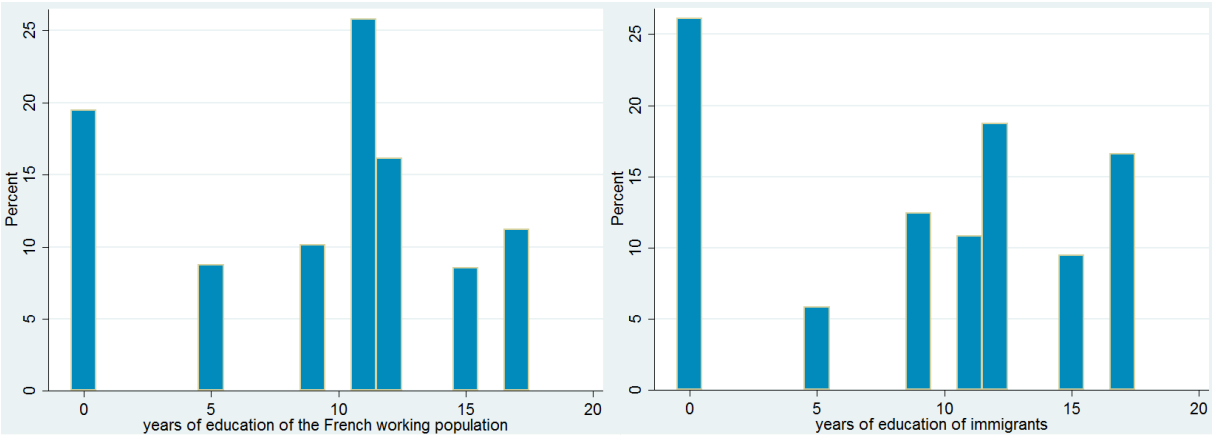
Table 2: Descriptive statistics

Variable	N	Min	Median	Max	Mean	Std. Dev.
General characteristics						
Age	6098	18	32	81	33.28	8.97
Time since arrival	6095	0	1	56	3.56	4.58
Number of children in the household	3097	0	1	8	1.45	1.11
Number of adults in the household	6004	1	2	10	2.19	1.02
Marital status (=1 if married)	6105	0	1	1	0.68	0.47
Years of education	5733	0	11	17	9.09	6.18

Variable	Men	Women	Overall
Age in 2010			
18 to 29	28.6	42.8	35.7
30 to 29	48.8	37.5	43.1
40 to 49	17.9	14.3	16.1
50 or more	4.7	5.4	5.0
Nationality			
Maghreb	32.7	32.4	32.6
Sub-Saharan Africa	30.4	25.0	27.7
Other African countries	8.1	7.3	7.7
Asia	18.8	18.5	18.7
Europe (excluding France) and CIS	4.8	9.3	7.1
America and Oceania	5.1	7.5	6.3
Time elapsed since arrival			
Less than two years	41.1	62.6	51.9
2 – 4 years	15.5	14.3	14.9
5 – 9 years	29.0	17.1	23.1
10 years or more	14.4	5.9	10.1
Migration motivation			
Professional reasons	19.2	4.7	11.9
Family reasons	58.1	80.3	69.2
Refugees	16.8	11.0	13.9
Others	5.8	4.0	4.9

Notes: Calculations based on the ELIPA database, first wave (2010). In the second part of the table, we report percentages.

Figure 1: Distribution of years of education in the French labour force and the immigrant's population



Source: Calculations from the ELIPA (first wave) and FQP databases.

Table 3: Distribution of the immigrant and the French working populations across socio-professional categories (in percent)

	Immigrant population	French working population
Farmers	0.0	3.8
Craftsmen, shopkeepers and business owners	2.6	5.0
Executives and higher intellectual occupations	3.1	13.1
Intermediate occupations	4.9	23.4
Employees (except those providing direct services to individuals)	12.5	24.2
Employees providing direct services to individuals	19.2	7.2
Skilled workers	19.2	14.8
Unskilled workers	25.6	8.4

Source: Calculations from the ELIPA (first wave) and FQP databases.

Table 4: Impact of language proficiency on over-education

	Dependent variable: overeduc				
	(1)	(2)	(3)	(4)	(5)
Self_report	.0473			.0393	.0390
	(.0423)			(.0498)	(.0498)
Not_fluent		.0713**			.0702**
		(.0316)			(.0315)
Oral			.0240	.0218	
			(.0384)	(.0383)	
Written			.0742**	.0743**	
			(.0361)	(.0361)	
Male	-.1925***	-.1651***	-.1669***	-.1718***	-.1699***
	(.0369)	(.0432)	(.0432)	(.0437)	(.0437)
Age	.0631***	.0583***	.0586***	.0583***	.0580***
	(.0037)	(.0042)	(.0042)	(.0042)	(.0042)
Exp	-.0405***	-.0376***	-.0375***	-.0377***	-.0378***
	(.0087)	(.0099)	(.0098)	(.0098)	(.0098)
Exp²	-.0004**	-.0004**	-.0004**	-.0004**	-.0004**
	(.0001)	(.0002)	(.0002)	(.0002)	(.0002)
Length	.0053	.0073	.0070	.0079	.0082
	(.0045)	(.0050)	(.0050)	(.0051)	(.0051)
Observations	1212	838	838	834	834
Adjusted R^2	0.3268	0.3320	0.3338	0.3337	0.3319

Notes: This table reports the results of probit regressions. The data are taken from the first wave of The Longitudinal Survey of the Integration of First-time Arrivals (ELIPA). Dummy dependent variable: individual education-occupation match category, taking the value 1 if the individual is over-educated. All regressions control for marital status, country of origin, household size and whether the migrant lives in IDF or not. The regressions present ROBUST standard errors. *, **, and *** indicate significance at the 10%, 5% and 1% level, respectively.

Table 5: Impact of networks on over-education

	Dependent variable: overeduc	
	(1)	(2)
Pre_mig	-.0349	
	(.0460)	
Post_mig	.0175	
	(.0554)	
French_network		-.0803**
		(.0316)
Foreign_network		.0014
		(.0436)
Male	-.1879***	-.1872***
	(.0372)	(.0371)
Age	.0637***	.0645***
	(.0037)	(.0037)
Exp	-.0408***	-.0437***
	(.0089)	(.0088)
Exp^2	-.0004**	-.0003**
	(.0001)	(.0001)
Length	.0048	.0036
	(.0045)	(.0045)
Observations	1198	1198
Adjusted R^2	0.3266	0.3303

Notes: This table reports the results of probit regressions. The data are taken from the first wave of The Longitudinal Survey of the Integration of First-time Arrivals (ELIPA). Dummy dependent variable: individual education-occupation match category, taking the value 1 if the individual is over-educated. All regressions control for marital status, country of origin, household size and whether the migrant lives in IDF or not. The regressions present ROBUST standard errors. *, **, and *** indicate significance at the 10%, 5% and 1% level, respectively.

Table 6: Robustness check: alternative definitions of over-education

	(A)	(B)	(C)
Oral	.0192	.0088	-.0021
	(.0382)	(.0130)	(.0026)
Written	.0708**	.0147	.0106***
	(.0358)	(.0120)	(.0050)
French_network	-.0789**	-.0184**	-.0077***
	(.0349)	(.0114)	(.0035)
Foreign_network	.0001	-.0074	-.0011
	(.0539)	(.0198)	(.0047)
Male	-.1610***	.0138	.0074***
	(.0437)	(.0106)	(.0037)
Age	.0594***	.0319***	.0070***
	(.0042)	(.0067)	(.0026)
Exp	-.0425***	-.0289***	-.0069***
	(.0098)	(.0077)	(.0028)
Exp^2	-.0003*	-.0001	-.0000
	(.0002)	(.0001)	(.0000)
Length	.0060	.0027	.0005
	(.0050)	(.0016)	(.0004)
Observations	825	825	825
Adjusted R^2	0.3391	0.5639	0.5167

Notes: This table reports the results of probit regressions. The data are taken from the first wave of The Longitudinal Survey of the Integration of First-time Arrivals (ELIPA). All regressions control for marital status, country of origin, household size and whether the migrant lives in IDF or not. Columns (A), (B) and (C) report the results from regressions on over-education, using the mode, the mean and the median values of educational attainment as the "required" level of education, respectively. The regressions present ROBUST standard errors. *, **, and *** indicate significance at the 10%, 5% and 1% level, respectively.

Table 7: Robustness check: alternative definitions of the language variables

	Dependent variable: overeduc			
	(A)	(A')	(B)	(B')
Not_fluent	.0641**			
	(.0318)			
Not_fluent2		.0796**		
		(.0308)		
Oral			.0192	
			(.0382)	
Written			.0708**	
			(.0358)	
Oral2				.0309
				(.0375)
Written2				.0589*
				(.0334)
French_network	-.0783**	-.0854**	-.0789**	-.0748**
	(.0358)	(.0317)	(.0349)	(.0358)
Foreign_network	-.0036	-.0082	.0001	.0041
	(.0544)	(.0431)	(.0539)	(.0547)
Male	-.1589***	-.1865***	-.1610***	-.1549***
	(.0438)	(.0371)	(.0437)	(.0434)
Age	.0592***	.0644***	.0594***	.0593***
	(.0042)	(.0037)	(.0042)	(.0042)
Exp	-.0426***	-.0438***	-.0425***	-.0425***
	(.0099)	(.0088)	(.0098)	(.0098)
Exp^2	-.0003*	-.0003**	-.0003*	-.0003
	(.0002)	(.0001)	(.0002)	(*.0002*)
Length	.0063	.0038	.0060	.0059
	(.0050)	(.0045)	(.0050)	(.0050)
Observations	825	1198	825	825
Adjusted R^2	0.3372	0.3341	0.3391	0.3382

Notes: This table reports the results of probit regressions. The data are taken from the first wave of The Longitudinal Survey of the Integration of First-time Arrivals (ELIPA). Dummy dependent variable: individual education-occupation match category, taking the value 1 if the individual is over-educated. All regressions control for marital status, country of origin, household size and whether the migrant lives in IDF or not. The regressions present ROBUST standard errors. *, **, and *** indicate significance at the 10%, 5% and 1% level, respectively.

Table 8: Robustness check: alternative econometric specification

	Dependent variable: overeduc			
	Probit (1)	Probit (2)	LPM (1)	LPM (2)
Not_fluent	.0641**		.0503**	
	(.0318)		(.0278)	
Oral		.0192		.0142
		(.0382)		(.0331)
Written		.0708**		.0528*
		(.0358)		(.0296)
French_network	-.0783**	-.0789**	-.0684**	-.0694**
	(.0358)	(.0349)	(.0339)	(.0336)
Foreign_network	-.0036	.0001	-.0055	-.0031
	(.0544)	(.0539)	(.0486)	(.0481)
Male	-.1589***	-.1610***	-.1308***	-.1310***
	(.0438)	(.0437)	(.047)	(.0346)
Age	.0592***	.0594***	.0443***	.0444***
	(.0042)	(.0042)	(.0026)	(.0027)
Exp	-.0426***	-.0425***	-.0370***	-.0368***
	(.0099)	(.0098)	(.0072)	(.0072)
Exp^2	-.0003*	-.0003*	-.0001	-.0001
	(.0002)	(.0002)	(.0001)	(.0001)
Length	.0063	.0060	.0042	.0040
	(.0050)	(.0050)	(.0043)	(.0043)
Intercept			-.2771	-.2890
			(.2904)	(.2843)
Observations	825	825	825	825
Adjusted R^2	0.3372	0.3391	0.2980	0.2981

Notes: This table reports the results of probit and linear probability model regressions. The data are taken from the first wave of The Longitudinal Survey of the Integration of First-time Arrivals (ELIPA). Dummy dependent variable: individual education-occupation match category, taking the value 1 if the individual is over-educated. All regressions control for marital status, country of origin, household size and whether the migrant lives in IDF or not. The regressions present ROBUST standard errors. *, **, and *** indicate significance at the 10%, 5% and 1% level, respectively.